Composition and Diversity of Endophytic Bacterial Community in Seeds of Upland Rice Resources from Different Origin Habitats in Yunnan Province of China

Zhishan Wang

University of Science and Technology

Yongqiang Zhu

Chinese National Human Genome Center at Shanghai

Ni Li

State Key Laboratory of Hybrid Rice (Hunan Hybrid Rice Research Center)

Biqi Sun

University of Science and Technology

Hai Liu

State Key Laboratory of Hybrid Rice (Hunan Hybrid Rice Research Center)

Huajun Zheng

Chinese National Human Genome Center at Shanghai

Weiping Wang

State Key Laboratory of Hybrid Rice (Hunan Hybrid Rice Research Center)

Yang Liu (✉ ly81150@163.com)

University of Science and Technology <https://orcid.org/0000-0003-3261-7243>

Research Article

Keywords: Yunnan Province of China, Upland rice seed, Endophytic bacteria, Diversity, Environmental factor

Posted Date: April 26th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-373041/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Upland rice has the characteristics of strong drought tolerance and wide adaptability. Cultivating upland rice with high yield and high quality can solve the contradiction between food shortage, water shortage, and population increase in countries all over the world, and is of great significance to the sustainable development of agriculture. This study aims to reveal the "core microbiota" of the endophytic bacteria in upland rice seeds in the Yunnan Province of China by examining their diversity and community structures. Through the correlation analysis with upland rice habitat environmental factors, the effects of climate and altitude on the structure and diversity of endophytic bacterial community in upland rice seeds were further revealed. In this study, high-throughput sequencing technology based on the Illumina Miseq platform was used to investigate the structure and diversity of endophytic bacterial communities using 12 upland rice variety seeds from different areas in Yunnan Province of China as materials. Here, 39 endophytic OTUs (0.68%) were found to coexist in all samples. At the phylum level, the first dominant phyla in the 12 seed samples were Proteobacteria (66.92–99.98%). At the genus level, *Pantoea* (9.75–99.24%), *Pseudomonas* (0.11–37.24%), *Curtobacterium* (0.01–19.90%), *Microbacterium* (0.01–14.95%), *Methylobacterium* (0.40–5.86%), *Agrobacterium* (0.01–4.53%), *Sphingomonas* (0.04–1.56%), *Aurantimonas* (0.01–1.45%) and *Rhodococcus* (0.11–1.09%), which represent the core microbiota in upland rice seeds, served as the dominant genera that coexisted in all the upland rice seeds tested. Environmental factors such as temperature, precipitation and altitude have great influences on the structure of endophytic bacterial community in upland rice seeds. This study is of great significance to explore the relationship between upland rice and its endophytic bacteria and to tap the resources of drought-tolerant bacteria to improve the yield of local upland rice.

Introduction

Plant endophyte is a kind of important microbial resource that lives in various tissues and organs of healthy plants at a certain stage or all stages without causing infection symptoms. In the process of long-term interaction, plants and endophytes have formed a symbiotic unit and become an important part of plant evolution. Through a large number of existing studies, it can be found that there are a large number of endophytic bacteria in the plants studied, which settle in the internal tissue of the host plant and form a series of mutually beneficial and symbiotic relationships (Hassani et al. 2018). Many studies have also found that the establishment of plant endophyte diversity and community structure is closely related to plant variety, genotype, growth environment and geographical location (Edwards et al. 2015; Santoyo et al. 2016; Rybakova et al. 2017; Liu et al. 2017, 2019, 2020; Walitang et al. 2018 a, b; Zhang et al. 2018b; Raj et al. 2019). In addition to the biological functions of endophyte, such as promoting seedling germination, promoting plant growth, increasing yield, enhancing resistance and inducing the synthesis of secondary metabolites, their natural products have potential in pharmaceutical, agricultural, industrial and other fields (Ansary et al. 2018; Mohamad et al. 2018; Rangjaroen et al. 2019; White et al. 2019).

Plant seed is not only the reproductive organ of the plant, the preserver and transmitter of plant genetic information, but also the adaptive strategy to ensure species reproduction under stress, which is of great

significance to plant reproduction and agricultural production (Steinbrecher and Leubner-Metzger 2017; Domergue et al. 2019). Studies have shown that plant seeds are rich in microbial resources, which are the transmission medium of microorganisms from maternal plants in offspring plants, and are responsible for passing beneficial microorganisms from generation to generation, and then achieve a direct or indirect impact on plant growth and development, health, quality, yield and functional components and other biological characteristics (Walitang et al. 2018a; Berg and Raaijmakers 2018; Li et al. 2019a; Liu et al. 2019). Endophytic bacteria are the most easily identified microorganisms by researchers. A large number of studies have found that the dominant endophytic bacteria in plant seeds are Proteobacteria, Actinobacteria, Firmicutes and Bacteroidetes phyla, and the dominant genera are *Pantoea*, *Bacillus*, *Pseudomonas*, *Sphingomonas*, *Microbacterium* and *Acinetobacter* (Truyens et al. 2015; Scott et al. 2018; Zhang et al. 2018b; Wang et al. 2020). In addition to the conventional culture-dependent method, 16S rDNA clone library technique and denaturing gradient gel electrophoresis (DGGE), which have been used to study endophytes in plant seeds, high-throughput sequencing has gradually become a new method to evaluate and analyze the diversity and community structure of endophytes in plant seeds (Hardoim et al. 2012; Jiang et al. 2013; Rybakova et al. 2017; Liu et al. 2017; Haimin et al. 2018). In the previous research, the undergraduate research group also used high-throughput sequencing technology to explore the endophytic bacterial diversity and community structure of 14 upland rice seeds collected in different areas of China (Wang et al. 2020). However, compared with other plant tissue microorganisms, research on plant seed endophytes is still very limited, in particular, there are few studies on the effects of environmental factors related to plant growth on the diversity and community structure of endophytic bacteria in their seeds.

Rice is the main food crop in China, and China is the country that produces the most rice in the world. However, the shortage of water resources in China has a great impact on rice cultivation, which forces researchers to step up their research on upland rice to deal with the problem. Upland rice is a kind of ecotype crop that is more tolerant and resistant to drought than rice. It is mainly planted in the Yellow-Huaihe River Basin of China and some places with insufficient water resources or uneven space and time. In-depth study of endophytic bacteria in upland rice seed is not only conducive to the improvement of its subsequent yield, but also of great help to explore the mechanism of drought tolerance. Based on the Illumina Miseq platform, the diversity and community structure of endophytic bacteria in 12 upland rice seeds from Yunnan Province of China were analyzed to explore the similarities and differences of endophytic bacteria in upland rice seeds growing in Yunnan Province of China. On this basis, combined with the temperature, humidity and altitude of each variety origin area, the effects of environmental factors of upland rice origin areas on the diversity and community structure of endophytic bacteria in upland rice seeds were studied. Besides, combined with our previous studies on the diversity and community structure of endophytic bacteria in upland rice seeds in different areas, we can further clarify the "core microbiota" of upland rice seeds.

Materials And Methods

The source of upland rice seeds

The 12 upland rice seed samples were provided by Hunan Hybrid Rice Research Center. The origin areas and related information of all samples are shown in Fig. 1 and Table 1. The samples were transferred to sterile bags, sealed and stored at 4 °C until used.

Sample surface sterilization and treatment

Three replicates of each sample were selected in this study. Firstly, the husks of each upland rice seed sample were removed by a small sheller. Then, under aseptic conditions the following operations were performed in the order listed: husked seeds were washed three times with prepared sterile water; 5 g of seeds were placed in a clean and sterile 50-mL tube containing 25 mL of phosphate buffer (per liter: 7.15 g of $\text{NaH}_2\text{PO}_4 \cdot 2\text{H}_2\text{O}$, 22.04 g of $\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$, 200 μL of Silwet L-77) (Zhang et al. 2018b; Liu et al. 2019), and the seeds were sonicated twice by an Ultrasonic Processor Scientz-IID sonicator (NingBo Scientz Biotechnology Co., Ltd., China) at low power (237.5 W; 950 W $\times$ 25%) in an ice bath for 5 min (alternating thirty 2-s bursts and thirty 2-s rests) (Zhang et al. 2018b; Liu et al. 2019). To validate that the surface was sterilized, sterile tweezers were used to press surface-sterilized seeds into LB medium (LUQIAO), and the samples were incubated at 30 °C for 72 h.

DNA extraction

Five gram of surface-sterilized upland rice seeds from each sample was frozen with liquid nitrogen and was quickly ground into a fine powder with a pre-cooled sterile mortar, and then the DNA was extracted using the FastDNA[®] SPIN Kit for Soil (MP Biomedicals, Solon, OH, USA) following the manufacturer's instructions of the Kit.

Amplicon library preparation and sequencing

All PCR amplifications were performed using TransStartFastPfu DNA Polymerase (TransGen, Beijing, China). For rice seeds, 799F (5'-AACAGGATTAGATACCCTG-3') and 1492R (5'-GGTTACCTTGTTACGACTT-3') was used for the first-round amplification (4-7bp barcode was added to the 5' primer of 968F and 1378R). Then the 750 bp fragment amplified from endophytic bacteria was cut and used as the template for the second-round amplification for the V6-V8 region (968F: 5'-AACGCGAAGAACCTTAC-3' and 1378R: 5'-CGGTGTGTACAAGGCCCGGGAACG-3'). All of the thermocycling steps are as follows: 5 min at 95 °C; 20 cycles of 45 sec at 95 °C, 30 sec at 57 °C, and 30 sec at 72 °C. The amplified products were purified and mixed in equivalent amounts. After using the TruePrep[®] DNA Library Prep Kit V2 for Illumina (Vazyme, China) to perform the library, all samples were sequenced on MiSeq (300-bp-paired-end sequencing) instrument using the MiSeq[®] Reagent Kit v3 (600 cycles) (Illumina).

Sequence data processing

The assembly of paired FASTQ files was performed by Mothur (version 1.39.0) (Schloss et al. 2011). Briefly, paired sequence reads were assembled after removing raw reads with ambiguous bases or low quality, such as read length < 50 bp, average Qscore < 25, or reads not matching the primer (pdiffs = 0)

and barcode (bdiffs = 0). The high-quality DNA sequences were aligned to SILVA reference database (V119) (Quast et al. 2013), and using chimera. uchime module to remove chimera sequences. Then the reads were classified and grouped into OTUs (operational taxonomic units) under the threshold of 97% identity.

Data statistics

Community richness, evenness and diversity analysis (Shannon, Simpson, ACE, Chao and Good's coverage) were performed using Mothur. Both PCoA and NMDS were analyzed based on the taylor matrix by mothur. The t-test (with 95% confidence intervals) was used to determine whether the means of evaluation indices were statistical difference, and p-value < 0.05 was considered as a significant standard. Taxonomy was assigned using the online software RDP classifier (Wang et al. 2007) at default parameter (80% threshold) based on the Ribosomal Database Project (Cole et al. 2009). Genera and family abundance differences between samples were analyzed by Metastats (White et al. 2009). Spearman correlation coefficient between two variables was calculated using the R command "cor.test". RDA analysis based on genus level was performed by "vegan" package in R. Average annual precipitation, average annual temperature and altitude were selected to be variable parameters.

Sequence accession numbers

The raw high-throughput sequencing data were submitted to the NCBI database with Accession number SRR13319808-SRR13319843 and BioProject number PRJNA688367.

Results

Diversity analysis of endophytic bacteria in upland rice seeds

According to the information of barcode and front-end primers, the quality control sequences were divided into 36 groups of sequence files, and a total of 2,089,709 high-quality sequences were obtained, with an average of 58,047 sequences per sample (Supplementary Table S1). Because of the large sample size, we use the added value of the repeated samples as the final calculation. The original diversity data are shown in Supplementary Table S2. According to the difference of distance between the sequences, 16S rRNA genes obtained were clustered into OTUs for species classification under the similarity level of 97%. A total of 5,704 OTUs were generated from all samples, with the number of OTUs in each sample ranging from 322 to 1,527 (Table 2). The statistical results found that 39 endophytic OTUs (0.68%) coexisted in the 12 sample seeds. And each sample contained its unique OTUs. The proportions of unique OTUs were 41.23%, 28.28%, 47.20%, 34.26%, 33.82%, 42.31%, 32.47%, 31.78%, 27.93%, 37.43%, 33.29% and 29.07% in sample 19H011, 19H012, 19H013, 19H014, 19H015, 19H017, 19H018, 19H019, 19H022, 19H024, 19H029 and 19H032, respectively, which indicated that the endophytic bacteria in the seeds of different upland rice varieties were different.

The diversity (α diversity) index of samples includes ACE, Chao, Shannon and Simpson values, in which ACE and Chao values are used for sample abundance assessment, and Shannon and Simpson values are used for sample diversity assessment. In general, there are some differences in diversity and abundance of each sample, among which sample 19H011, 19H015, 19H017, 19H018 and 19H024 have significant differences in ACE, Chao, Shannon and Simpson values from other samples (Table 2). The values of ACE, Chao and Shannon of sample 19H017 were significantly higher than those of other samples, and the values of sample 19H013 were the lowest, which indicated that the diversity and richness of endophytic bacteria in sample 19H017 were the highest among all samples, and that of sample 19H013 was the lowest. The above results are consistent with the results of Rarefaction curve and Rank abundance curve of all samples in the case of OUT = 0.03 (Supplementary Fig. S1 and Fig. S2).

Bacterial endophytes community compositions and structures of upland rice seeds

The endophytic bacterial community composition of 12 upland rice seed samples from different areas in Yunnan Province of China is shown in Fig. 2 at the phylum level. The endophytic bacterial community structure of 12 samples had low diversity at the phylum level, mainly including Proteobacteria and Actinobacteria. The relative abundance of Proteobacteria was relatively high, and the relative abundance of different samples was between 66.92% and 99.98%, which was the main group of bacteria. And followed by Actinomycetes with abundance ranging from 0.01% to 32.21%. At the genus level, 148 genera were covered by endophytic bacteria in all upland rice seed samples, and the main bacteria with high relative abundance were *Pantoea*, *Pseudomonas*, *Curtobacterium*, *Microbacterium*, *Methylobacterium*, *Agrobacterium*, *Sphingomonas*, *Aurantimonas* and *Rhodococcus*, and the proportion ranges from 9.75-99.24%, 0.11-37.24%, 0.01-19.90%, 0.01-14.95%, 0.40-5.86%, 0.01-4.53%, 0.04-1.56%, 0.01-1.45% and 0.11-1.09%, respectively (Fig. 3). Table 3 listed in detail the dominant genera and proportion of each upland rice seed sample, which indicated that there were significant differences in the abundance distribution of endophytic bacteria in different upland rice seed samples. The classification of samples at the level of 97% sequence similarity (genus, top10) in Fig. 4 also showed that the abundance distribution of endophytic bacteria in each seed sample was different at the genus level.

To explore the differences in the community structure of endophytic bacteria in the 12 upland rice seed samples, PCoA (Principal Co-ordinates Analysis) and NMDS (Non-metric Multidimensional Scaling) methods were used to draw the two-dimensional distribution diagram of seed samples (Fig. 5 and Fig. 6). The distance of each sample in the two-dimensional diagram can reflect the degree of community structure similarity, and the closer the distance between sample points is, the more similar the community structure is. The results of PCoA and NMDS showed that the distance between sample 19H019 and 19H029 was close, and the distance between sample 19H012, 19H015, 19H022 and 19H024 was close, which indicated that the endophytic bacterial community structure of sample 19H019 and 19H029 was similar, and the endophytic bacterial community structure of sample 19H012, 19H015, 19H022 and 19H024 was similar.

Analysis of environmental factors affecting the community structure and diversity of endophytic bacteria in upland rice seeds

In order to further explore the impact of environmental factors of origin areas on the endophytic bacterial community structure and diversity of the sample seeds, we inquired about the temperature, precipitation and altitude of the corresponding areas in Yunnan Province of China. The specific data are shown in Table 4 and Fig. 1B. Then we showed the relationship between environmental factors and sample distribution and the main dominant bacteria by RDA (Redundancy Analysis), taking the average annual precipitation, average annual temperature and altitude as variables (Fig. 7). The effect of environmental factors on endophytic bacteria in seeds in the RDA diagram is mainly characterized by the length of environmental factors, while the influence degree of environmental factors on each strain is reflected by the cosine value of the angle. Temperature, precipitation and altitude have great effects on endophytic bacteria in upland rice seeds, among which precipitation and altitude are the main influencing factors (Fig. 7). There was a significant positive correlation between the main dominant bacteria *Pantoea* and precipitation, temperature and altitude, and the correlation between *Pantoea* and altitude was the strongest. However, there was a negative correlation between other dominant bacteria and environmental factors. The proportion of other bacteria is low and no further analysis is made.

Discussion

The food issue is a major issue related to the national economy and people's livelihood, and food security is an important part of national security. As we all know, the three major food crops in the world are wheat, rice and corn, among which wheat is the food crop with the largest sown area, the largest yield, and the most widely distributed in the world, while rice ranks second. According to statistics, rice is cultivated in 122 countries in the world, with a perennial cultivation area of 140–150 million hectares, which is widely distributed. However, droughts caused by persistent climate instability and unpredictable rainfall patterns have had a significant impact on rice cultivation, especially in sub-Saharan Africa and Southeast Asian countries, which makes upland rice cultivation a breakthrough to solve this problem (Khan et al. 2020). At present, the research on upland rice is mainly focused on its drought tolerance and genetic improvement (Xia et al. 2019; Abdirad et al. 2020; Luo et al. 2020b; Uddin and Fukuta 2020). Studies on the root microbial diversity of upland rice and its effects on the growth and drought tolerance of upland rice have also been reported (Pang et al. 2020). Our research group also explored the diversity and community structure of endophytic bacteria in upland rice seeds in different regions of China and preliminarily revealed the core microbiota of endophytic bacteria in upland rice seeds (Wang et al. 2020).

Under the circumstances of global population expansion, climate warming and water shortage, the development and utilization of upland rice with drought tolerance, barren tolerance and wide adaptability is of great significance to ensure global food security. In this research, We obtained 2,089,709 effective sequences of endophytic bacteria by Illumina Miseq platform sequencing and divided them into 5,704 OTUs, which could reflect the composition of endophytic bacteria in all samples. The results showed that the coexistence OTUs of the 12 upland rice seed samples were 39 (0.68%), which indicated that the main

microbial groups of these upland rice seed samples were similar. Proteobacteria, Firmicutes, Bacteroidetes and Actinobacteria are the main dominant bacteria often reported in plant seeds (Truyens et al. 2015; Scott et al. 2018; Zhang et al. 2018b; Wang et al. 2020), while Proteobacteria (66.92%-99.98%) and Actinomycetes (0.01%-32.21%) also represented the main dominant phyla of endogenous bacteria in upland rice seeds in this study.

Overall, this study found that the main endophytic bacterial groups and community structure of the 12 upland rice seed samples in Yunnan Province of China were similar. The overlapping and non-overlapping areas of the Venn diagram are usually used to represent the shared OTUs and unique OTUs existing in each group of samples or each sample (Vandenkoornhuyse et al. 2015), but due to the large number of samples, the Venn diagram could not be drawn. Through the statistics of Venn diagram data (Supplementary Table S3), it is found that the shared OTUs of all samples was 0.68%, of which the abundance accounted for > 2.55% of the total OTUs in each seed sample. The primary shared genera were *Pantoea*, *Pseudomonas*, *Curtobacterium*, *Microbacterium*, *Methylobacterium*, *Agrobacterium*, *Sphingomonas*, *Aurantimonas* and *Rhodococcus*. *Pantoea* (9.75–99.24%) was the first dominant genus shared by all 12 seed samples, while other shared genera were represented by *Pseudomonas* (0.11–37.24%), *Curtobacterium* (0.01–19.90%), *Microbacterium* (0.01–14.95%), *Methylobacterium* (0.40–5.86%), *Agrobacterium* (0.01–4.53%), *Sphingomonas* (0.04–1.56%), *Aurantimonas* (0.01–1.45%) and *Rhodococcus* (0.11–1.09%). Although most species of *Pantoea* are usually well-known plant pathogens, in our previous studies on the endophytic bacteria of upland rice seeds collected in different regions, we also found that *Pantoea* was the first dominant genus in all samples (Wang et al. 2020). Besides, a large number of studies have also reported that *Pantoea* is the endophytic dominant genus in rice, barley, maize, rape, hemp and *Salvia miltiorrhiza* seeds (Granér et al. 2003; Liu et al. 2017, 2019; Shahzad et al. 2017; Walitang et al. 2017; Chen et al. 2018; Rahman et al. 2018; Scott et al. 2018; Zhang et al. 2018b; Raj et al. 2019). As a dominant endophyte, potentially pathogenic bacteria *Pantoea* exists in healthy upland rice and other plant seeds, but does not show disease symptoms, indicating that it should have other effects on plants. It has been proved that some species of *Pantoea* do play an important role in promoting plant growth. Luziatelli et al. (2020) have found that *Pantoeaagglomerans* C1 isolated from lettuce not only has biological functions such as producing indole-3-acetic acid (IAA), dissolving phosphate and inhibiting plant pathogens, but also has high biotechnology potential as a plant growth-promoting bacteria in heavy metal contaminated soil. Another study has also found that endophytic bacteria *Pantoeaagglomerans* can well alleviate abiotic stresses such as drought and salt in durum wheat (Cherif-Silini et al. 2019). Besides, a large number of reports show that *Pantoea* is of great help to plant growth, yield increase and disease resistance (Megías et al. 2016; Lumactud and Fulthorpe 2018; Zhang et al. 2018a).

In addition to *Pantoea*, the main genus of shared endophytic bacteria, some dominant genera, including *Pseudomonas*, *Curtobacterium*, *Microbacterium*, *Methylobacterium*, *Agrobacterium*, *Sphingomonas*, *Aurantimonas* and *Rhodococcus*, are also shared by all 12 upland rice seeds, which may be due to the conservation of plant seeds. Interestingly, this result is basically consistent with our previous studies on endophytic bacteria in upland rice seeds. Interestingly, this result is basically consistent with our previous

studies on endophytic bacteria in upland rice seeds, which once again strongly indicates that these bacteria are indeed the core endophytic bacterial communities in upland rice seeds. *Pseudomonas* is usually the dominant genus in the seeds of barley (Rahman et al. 2018), rice (Walitang et al. 2017; Liu et al. 2019), maize (Liu et al. 2012; Liu et al. 2020), rape (Granér et al. 2003), sugar beet (Kanivets and Pishchur 2001), peanuts (Sobolev et al. 2013), browntop millet (Verma and White 2018), Marama beans (Chimwamurombe et al. 2016), hemp (Scott et al. 2018), tobacco (Li et al. 2019b) and *Salvia miltiorrhiza* (Chen et al. 2018). Besides, *Curtobacterium*, *Microbacterium*, *Methylobacterium*, *Agrobacterium*, *Sphingomonas* and *Aurantimonas* are often found in the seeds of these plants as endophytic dominant genus. It is also reported that many species of these bacteria have important contributions in promoting plant growth and improving plant nitrogen fixation, resistance to harmful bacteria, cold resistance, drought resistance and tolerance to heavy metals. For example, Kumawat et al. (2019) found that the synergism of *Pseudomonas aeruginosa* and *Bradyrhizobium* sp. can improve plant growth, nutrient acquisition and soil health in soybean. Another study found that endophytic bacteria *Pseudomonas stutzeri* A15 isolated from rice had strong nitrogen fixation ability, which was significantly better than that of chemical nitrogen fertilizer after inoculation of rice (Pham et al. 2017). The researchers also found that endophytic bacteria *Curtobacterium*, *Microbacterium* and *Methylobacterium* carried in plant tissues generally can promote host plants to produce auxin, dissolve phosphate and inhibit *Fusarium* and other fungal pathogens (Sánchez-López et al. 2018; Verma and White 2018; Passari et al. 2019). Of course, for upland rice, the ability of endophytic dominant bacteria in their seeds to withstand the abiotic stress of drought is undoubtedly the most important. According to the existing research, it can be found that *Pseudomonas*, *Microbacterium*, *Methylobacterium* and *Sphingomonas* isolated from plants do play a great role in improving the drought tolerance and drought resistance of plants (Wang et al. 2014; Egamberdieva et al. 2015; García-Fontana et al. 2020; Luo et al. 2020a; Zhang et al. 2020). Therefore, it is meaningful to explore the drought tolerance and mechanism of drought tolerance from the perspective of endophytic bacteria in upland rice seeds. It is worth mentioning that different from other genera of endophytic dominant bacteria, *Rhodococcus* is rarely reported as a dominant genus in plant endophytes. The ability of *Rhodococcus* to degrade a variety of toxic chemicals and produce bioactive substances is the most frequently reported, and more and more attention has been paid to it. Reports on the *Rhodococcus* also show that some of its species not only have the function of producing cytokinin but also can colonize the root system and increase the plant biomass in polychlorinated biphenyl (PCBs) contaminated soil (Jameson 2019; Vergani et al. 2019).

Through this study, we also found that although there were significant differences in endophytic bacterial diversity, abundance and community structure in some of the 12 upland rice samples from origin areas of Yunnan Province of China, the difference of some of the samples was very small. The results of PCoA and NMDS showed that all samples could be separated in PC1- PC2 or NMDS1-NMDS2 coordinate system (Fig. 5 and Fig. 6), and there were differences among different upland rice seeds by comparing the diversity index of different upland rice varieties (Table 2). This further shows that the differences in varieties and genotypes of upland rice do have a certain effect on the diversity and community structure of endophytic bacteria in upland rice seeds. A large number of studies on the community structure and

diversity of endophytic bacteria in plant seeds have also found that although there are coexisting endophytic dominant bacterial groups among different varieties of plant seeds, at the same time, they have significant differences in endophytic bacterial community structure and diversity (Xu et al. 2014; Rybakova et al. 2017; Liu et al. 2017, 2019, 2020; López et al. 2018; Walitang et al. 2018 ab; Raj et al. 2019). Although some studies have found that the endophytic bacterial community structure and diversity among different varieties of seeds are very similar, there are still some differences in the abundance of bacterial groups (Zhang et al. 2018b; Wang et al. 2020). However, there are no significant differences in endophytic bacterial diversity and community structure in some upland rice seeds, and they are particularly close in PCoA and NMDS results (Fig. 5 and Fig. 6). Therefore, the diversity and community structure of endophytic bacteria in upland rice seeds should be affected not only by variety differences, but also by other factors, such as environmental factors of origin areas, structure and chemical composition of samples (Torres-Cortés et al. 2018; Girsowicz et al. 2019; Kuźniar et al. 2020).

In order to further explore the influencing factors of endophytic bacterial community structure and diversity in upland rice seed samples, we investigated and compared the environmental factors such as temperature, precipitation and altitude of origin areas in Yunnan Province of China. There were differences in these environmental factors among the origin areas, and through RDA analysis, it was further found that temperature, precipitation and altitude had great effects on endophytic bacteria. Among them, the altitude at the origin areas had the greatest influence on the main dominant bacteria *Pantoea* in upland rice seeds. Thus it can be seen that the community structure and diversity of endophytic bacteria in upland rice seeds are affected not only by upland rice varieties and genotypes, but also by environmental factors such as temperature, precipitation and altitude. We can also conclude that the community structure and composition of endophytic bacteria in upland rice seeds are caused by upland rice varieties, genotypes and environment, rather than by a single factor.

The actual living state of plants in nature is the state of microorganisms and plants, and plant breeding is the cultivation of symbiotes between plants and microorganisms (Wang et al. 2015). Upland rice has the characteristics of drought resistance and drought tolerance, and the symbiotic microorganisms should also have corresponding drought tolerance characteristics to adapt to the local environment. Using high-throughput sequencing technology to explore the community structure and diversity of endophytic bacteria in upland rice seeds from origin areas of Yunnan Province of China is of great significance for the subsequent excavation of drought-tolerant bacteria resources and the improvement of local upland rice yield. At the same time, the mechanism of drought tolerance at the microbial level of upland rice can be reflected by comparing the microbial differences between upland rice and rice, and this study lays a foundation for this.

Conclusion

Exploring the endophytic microbial community structure and diversity of upland rice seeds is the basis for understanding the synergistic effect of endophytic bacteria in upland rice and the new functions and new substances produced by the synergy. *Pantoea*, *Pseudomonas*, *Curtobacterium*, *Microbacterium*,

Methylobacterium, *Agrobacterium*, *Sphingomonas*, *Aurantimonas* and *Rhodococcus* served as major core endophytic bacteria in twelve upland rice seed samples in this study. Overall, there were some differences in endophytic bacterial community structure and diversity among all seed samples, but the differences in some samples were not significant. The differences of endophytic bacterial community structure and diversity in upland rice seeds are not only related to the varieties of the samples themselves, but also affected by local temperature, precipitation, altitude and other environmental factors.

Abbreviations

OTU: Optical Transform Unit; DGGE: Denaturing Gradient Gel Electrophoresis; PCoA: Principal Coordinates Analysis; NMDS: Non-metric Multidimensional Scaling; RDA: Redundancy Analysis

Declarations

Acknowledgments

We would like to thank the research team of Prof. Huajun Zheng of Shanghai-MOST Key Laboratory of Health and Disease Genomics of Chinese National Human Genome Center at Shanghai for its platform and technical support in high-throughput sequencing and bioinformatics analysis.

Authors' Contributions

Wang ZS and Zhu YQ designed and participated in all experimental procedures, performed data analysis, and drafted the manuscript. Sun BQ participated in the samples collection and preparation. Li N and Liu H participated in the plant samples cultivation. Zheng HJ, Wang WP and Liu Y supervised the study and critically revised the manuscript. All authors read and approved the final manuscript.

Funding

The research was supported by the Fundamental Research Funds for the Central Universities (No. FRF-TP-20-044A2; FRF-BR-19-003B; FRF-BR-20-03B) and the Open Research Fund of State Key Laboratory of Hybrid Rice (Hunan Hybrid Rice Research Center) (No. 2017KF04).

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Compliance with ethical standards

This article does not contain any studies with human participants or animals performed by any of the authors.

Conflict of interest

The authors declare that they have no competing interests.

References

1. Abdirad S, Majd A, Irian S, Hadidi N, Hosseini Salekdeh G (2020) Differential adaptation strategies to different levels of soil water deficit in two upland and lowland genotypes of rice: a physiological and metabolic approach. *J Sci Food Agric* 100(4):1458–1469
2. Ansary WR, Prince FRK, Haque E, Sultana F, West HM, Rahman M, Mondol AM, Akanda AM, Rahman M, Clarke ML, Islam T (2018) Endophytic *Bacillus* spp. from medicinal plants inhibit mycelial growth of *Sclerotinia sclerotiorum* and promote plant growth. *Z Naturforsch C J Biosci* 73(5–6):247–256
3. Berg G, Raaijmakers JM (2018) Saving seed microbiomes. *ISME J* 12(5):1167–1170
4. Chen H, Wu H, Yan B, Zhao H, Liu F, Zhang H, Sheng Q, Miao F, Liang Z (2018) Core microbiome of medicinal plant *Salvia miltiorrhiza* seed: a rich reservoir of beneficial microbes for secondary metabolism? *Int J Mol Sci* 19(3):672
5. Cherif-Silini H, Thissera B, Bouket AC, Saadaoui N, Silini A, Eshelli M, Alenezi FN, Vallat A, Luptakova L, Yahiaoui B, Cherrad S, Vacher S, Rateb ME, Belbahri L (2019) Durum wheat stress tolerance induced by endophyte *Pantoea agglomerans* with genes contributing to plant functions and secondary metabolite arsenal. *Int J Mol Sci* 20(16):3989
6. Chimwamurombe PM, Grönemeyer JL, Reinhold-Hurek B (2016) Isolation and characterization of culturable seed-associated bacterial endophytes from gnotobiotically grown Maramba bean seedlings. *FEMS Microbiol Ecol* 92(6):fiw083
7. Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ, Kulam-Syed-Mohideen AS, McGarrell DM, Marsh T, Garrity GM, Tiedje JM (2009) The ribosomal database project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res* 37(2):D141–D145
8. Domergue JB, Abadie C, Limami A, Way D, Tcherkez G (2019) Seed quality and carbon primary metabolism. *Plant Cell Environ* 42(10):2776–2788
9. Edwards J, Johnson C, Santosmedellín C, Lurie E, Podishetty NK, Bhatnagar S, Eisen JA, Sundaresan V (2015) Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc Natl Acad Sci* 112:911–920
10. Egamberdieva D, Wirth S, Alqarawi AA, Abd Allah EF (2015) Salt tolerant *Methylobacterium mesophilicum* showed viable colonization abilities in the plant rhizosphere. *Saudi*

11. García-Fontana C, Vilchez JI, Manzanera M (2020) Proteome comparison between natural desiccation-tolerant plants and drought-protected *Caspicum annuum* plants by *Microbacterium* sp. 3J1. Front Microbiol 11:1537
12. Girsowicz R, Moroenyane I, Steinberger Y (2019) Bacterial seed endophyte community of annual plants modulated by plant photosynthetic pathways. Microbiol Res 223–225:58–62
13. Granér G, Persson P, Meijer J, Alström S (2003) A study on microbial diversity in different cultivars of *Brassica napus* in relation to its wilt pathogen, *Verticillium longisporum*. FEMS Microbiol Lett 224(2):269–276
14. Haimin C, Hongxia W, Bin Y, Hongguang Z, Fenghua L, Haihua Z, Qing S, Fang M, Zongsuo L (2018) Core microbiome of medicinal plant *Salvia miltiorrhiza* seed: a rich reservoir of beneficial microbes for secondary metabolism? Int J Mol Sci 19(3):672
15. Hassani MA, Durán P, Hacquard S (2018) Microbial interactions within the plant holobiont. Microbiome 6(1):58
16. Hardoim PR, Hardoim CC, van Overbeek LS, van Elsas JD (2012) Dynamics of seed-borne rice endophytes on early plant growth stages. PLoS ONE 7(2):e30438
17. Jameson PE (2019) Virulent *Rhodococcus fascians* produce unique methylated cytokinins. Plants (Basel) 8(12):582
18. Jiang XY, Gao JS, Xu FH, Cao YH, Tang X, Zhang XX (2013) Diversity of endophytic bacteria in rice seeds and their secretion of indole acetic acid (in Chinese). Acta Microbiol Sin 53:269–275
19. Kanivets VI, Pishchur IN (2001) Bacterial microflora on disinfected sugar beet seeds. Microbiol 70(3):316–318
20. Khan MIR, Palakolanu SR, Chopra P, Rajurkar AB, Gupta R, Iqbal N, Maheshwari C (2020) Improving drought tolerance in rice: Ensuring food security through multi-dimensional approaches. Physiol Plant. <https://doi.org/10.1111/ppl.13223>
21. Kumawat KC, Sharma P, Sirari A, Singh I, Gill BS, Singh U, Saharan K (2019) Synergism of *Pseudomonas aeruginosa* (LSE-2) nodule endophyte with *Bradyrhizobium* sp. (LSBR-3) for improving plant growth, nutrient acquisition and soil health in soybean. World J Microbiol Biotechnol 35(3):47
22. Kuźniar A, Włodarczyk K, Grządziel J, Woźniak M, Furtak K, Gałązka A, Dziadczyk E, Skórzyńska-Polit E, Wolińska A (2020) New insight into the composition of wheat seed microbiota. Int J Mol Sci 21(13):4634
23. Li L, Zhang Z, Pan S, Li L, Li X (2019a) Characterization and metabolism effect of seed endophytic bacteria associated with peanut grown in south China. Front Microbiol 10:2659
24. Li J, Zheng B, Hu R, Liu Y, Jing Y, Xiao Y, Sun M, Chen W, Zhou Q (2019b) *Pseudomonas* species isolated from tobacco seed promote root growth and reduce lead contents in *Nicotiana tabacum* K326. Can J Microbiol 65(3):214–223

25. Liu Y, Zuo S, Xu L, Zou Y, Song W (2012) Study on diversity of endophytic bacterial communities in seeds of hybrid maize and their parental lines. *Arch Microbiol* 194(12):1001–1012
26. Liu Y, Wang R, Li Y, Gao Y, Chen C, Qiu C, Bai F, Xu T, Zhang X, Dai W, Zhao J, Cheng C (2017) High-throughput sequencing-based analysis of the composition and diversity of endophytic bacterial community in seeds of "Beijing" hybrid maize planted in China. *Plant Growth Regul* 81:317–324
27. Liu Y, Xu PP, Yang FZ, Li M, Yan H, Li N, Zhang XX, Wang WP (2019) Composition and diversity of endophytic bacterial community in seeds of super hybrid rice 'Shenliangyou 5814' (*Oryza sativa* L.) and its parental lines. *Plant Growth Regul* 87:257–266
28. Liu Y, Yan H, Zhang X, Zhang R, Li M, Xu T, Yang F, Zheng H, Zhao J (2020) Investigating the endophytic bacterial diversity and community structures in seeds of genetically related maize (*Zea mays* L.) genotypes. *3 Biotech* 10(1):27
29. López SMY, Pastorino GN, Franco MEE, Medina R, Lucentini CG, Saparrat MCN, Balatti PA (2018) Microbial endophytes that live within the seeds of two tomato hybrids cultivated in argentina. *Agronomy* 8:136
30. Lumactud R, Fulthorpe RR (2018) Endophytic bacterial community structure and function of herbaceous plants from petroleum hydrocarbon contaminated and non-contaminated sites. *Front Microbiol* 9:1926
31. Luo Y, Zhou M, Zhao Q, Wang F, Gao J, Sheng H, An L (2020a) Complete genome sequence of *Sphingomonas* sp. Cra20, a drought resistant and plant growth promoting rhizobacteria. *Genomics* 112(5):3648–3657
32. Luo Z, Xiong J, Xia H, Ma X, Gao M, Wang L, Liu G, Yu X, Luo L (2020b) Transcriptomic divergence between upland and lowland ecotypes contributes to rice adaptation to a drought-prone agroecosystem. *Evol Appl* 13(9):2484–2496
33. Luziatelli F, Ficca AG, Cardarelli M, Melini F, Cavalieri A, Ruzzi M (2020) Genome sequencing of *Pantoea agglomerans* C1 provides insights into molecular and genetic mechanisms of plant growth-promotion and tolerance to heavy metals. *Microorganisms* 8(2):153
34. Megías E, Megías M, Ollero FJ, Hungria M (2016) Draft genome sequence of *Pantoeaananatis* strain AMG521, a rice plant growth-promoting bacterial endophyte isolated from the Guadalquivir marshes in southern Spain. *Genome Announc* 4(1):e01681–e01615
35. Mohamad OAA, Li L, Ma JB, Hatab S, Xu L, Guo JW, Rasulov BA, Liu YH, Hedlund BP, Li WJ (2018) Evaluation of the antimicrobial activity of endophytic bacterial populations from Chinese traditional medicinal plant licorice and characterization of the bioactive secondary metabolites produced by *Bacillus atrophaeus* against *Verticillium dahliae*. *Front Microbiol* 9:924
36. Pang Z, Zhao Y, Xu P, Yu D (2020) Microbial diversity of upland rice roots and their influence on rice growth and drought tolerance. *Microorganisms* 8(9):1329
37. Passari AK, Rajput V, Zothanpuia, Priya LPM, Dharne M, Dastager S, Mathew OK, Hashem A, Abd Allah EF, Singh BP (2019) Draft genome sequence of plant growth-promoting endophytic

Microbacteriumhydrothermale BPSAC84, isolated from the medicinal plant *Mirabilis jalapa*.

Microbiol ResourAnnounc 8(22):e00406–e00419

38. Pham VT, Rediers H, Ghequire MG, Nguyen HH, De Mot R, Vanderleyden J, Spaepen S (2017) The plant growth-promoting effect of the nitrogen-fixing endophyte *Pseudomonas stutzeri* A15. Arch Microbiol 199(3):513–517
39. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glockner FO (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41:D590–D596
40. Rahman MM, Flory E, Koyro HW, Abideen Z, Schikora A, Suarez C, Schnell S, Cardinale M (2018) Consistent associations with beneficial bacteria in the seed endosphere of barley (*Hordeum vulgare* L.). Syst Appl Microbiol 41(4):386–398
41. Raj G, Shadab M, Deka S, Das M, Baruah J, Bharali R, Talukdar NC (2019) Seed interior microbiome of rice genotypes indigenous to three agroecosystems of Indo-Burma biodiversity hotspot. BMC Genom 20(1):924
42. Rangjaroen C, Lumyong S, Sloan WT, Sungthong R (2019) Herbicide-tolerant endophytic bacteria of rice plants as the biopriming agents for fertility recovery and disease suppression of unhealthy rice seeds. BMC Plant Biol 19(1):580
43. Rybakova D, Mancinelli R, Wikström M, Birch-Jensen AS, Postma J, Ehlers RU, Goertz S, Berg G (2017) The structure of the *Brassica napus* seed microbiome is cultivar-dependent and affects the interactions of symbionts and pathogens. Microbiome 5(1):104
44. Sánchez-López AS, Pintelon I, Stevens V, Imperato V, Timmermans JP, González-Chávez C, Carrillo-González R, Van Hamme J, Vangronsveld J, Thijs S (2018) Seed endophyte microbiome of *Crotalaria pumila* unpeeled: identification of plant-beneficial *Methylobacteria*. Int J Mol Sci 19(1):291
45. Santoyo G, Moreno-Hagelsieb G, Orozco-MosquedaMdel C, Glick BR (2016) Plant growth-promoting bacterial endophytes. Microbiol Res 183:92–99
46. Schloss PD, Gevers D, Westcott SL (2011) Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. PLoS ONE 6(12):e27310
47. Scott M, Rani M, Samsatly J, Charron JB, Jabaji S (2018) Endophytes of industrial hemp (*Cannabis sativa* L.) cultivars: identification of culturable bacteria and fungi in leaves, petioles, and seeds. Can J Microbiol 64(10):664–680
48. Shahzad R, Waqas M, Khan AL, Al-Hosni K, Kang SM, Seo CW, Lee IJ (2017) Indoleacetic acid production and plant growth promoting potential of bacterial endophytes isolated from rice (*Oryza sativa* L.) seeds. Acta Biol Hung 68(2):175–186
49. Sobolev VS, Orner VA, Arias RS (2013) Distribution of bacterial endophytes in peanut seeds obtained from axenic and control plant material under field conditions. Plant Soil 371:367–376
50. Steinbrecher T, Leubner-Metzger G (2017) The biomechanics of seed germination. J Exp Bot 68(4):765–783

51. Truyens S, Weyens N, Cuypers A, Vangronsveld J (2015) Bacterial seed endophytes: genera, vertical transmission and interaction with plants. *Environ Microbiol Rep* 7:40–50
52. Torres-Cortés G, Bonneau S, Bouchez O, Genthon C, Briand M, Jacques MA, Barret M (2018) Functional microbial features driving community assembly during seed germination and emergence. *Front Plant Sci* 9:902
53. Uddin MN, Fukuta Y (2020) A region on chromosome 7 related to differentiation of rice (*Oryza sativa* L.) between lowland and upland ecotypes. *Front Plant Sci* 11:1135
54. Vandenkoornhuyse P, Quaiser A, Duhamel M, Le Van A, Dufresne A (2015) The importance of the microbiome of the plant holobiont. *New Phytol* 206:1196–1206
55. Vergani L, Mapelli F, Suman J, Cajthaml T, Uhlik O, Borin S (2019) Novel PCB-degrading *Rhodococcus* strains able to promote plant growth for assisted rhizoremediation of historically polluted soils. *PLoS One* 14(8):e0221253
56. Verma SK, White JF (2018) Indigenous endophytic seed bacteria promote seedling development and defend against fungal disease in browntop millet (*Urochloa ramosa* L.). *J Appl Microbiol* 124(3):764–778
57. Walitang DI, Kim K, Madhaiyan M, Kim YK, Kang Y, Sa T (2017) Characterizing endophytic competence and plant growth promotion of bacterial endophytes inhabiting the seed endosphere of Rice. *BMC Microbiol* 17(1):209
58. Walitang DI, Kim CG, Jeon S, Kang Y, Sa T (2018a) Conservation and transmission of seed bacterial endophytes across generations following crossbreeding and repeated inbreeding of rice at different geographic locations. *Microbiologyopen* 8(3):e00662
59. Walitang DI, Kim CG, Kim K, Kang Y, Kim YK, Sa T (2018b) The influence of host genotype and salt stress on the seed endophytic community of salt-sensitive and salt-tolerant rice cultivars. *BMC Plant Biol* 18(1):51
60. Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 73(16):5261–5267
61. Wang S, Ouyang L, Ju X, Zhang L, Zhang Q, Li Y (2014) Survey of plant drought-resistance promoting bacteria from *Populus euphratica* tree living in arid area. *Indian J Microbiol* 54(4):419–426
62. Wang ZW, Ji YL, Chen YG (2015) Studies and biological significances of plant endophytes. *Microbiology China* 42(2):349–363
63. Wang Z, Zhu Y, Jing R, Wu X, Li N, Liu H, Zhang X, Wang W, Liu Y (2020) High-throughput sequencing-based analysis of the composition and diversity of endophytic bacterial community in seeds of upland rice. *Arch Microbiol*. [. org/10.1007/s00203-020-02058-9](https://doi.org/10.1007/s00203-020-02058-9)
64. White JR, Nagarajan N, Pop M (2009) Statistical methods for detecting differentially abundant features in clinical metagenomic samples. *PLoS Comput Biol* 5(4):e1000352
65. White JF, Kingsley KL, Zhang Q, Verma R, Obi N, Dvinskikh S, Elmore MT, Verma SK, Gond SK, Kowalski KP (2019) Review: Endophytic microbes and their potential applications in crop

management. Pest Manag Sci 75(10):2558–2565

66. Xia H, Luo Z, Xiong J, Ma X, Lou Q, Wei H, Qiu J, Yang H, Liu G, Fan L, Chen L, Luo L (2019) Bi-directional selection in upland rice leads to its adaptive differentiation from lowland rice in drought resistance and productivity. Mol Plant 12(2):170–184
67. Xu M, Sheng J, Chen L, Men Y, Gan L, Guo S, Shen L (2014) Bacterial community compositions of tomato (*Lycopersicum esculentum* Mill.) seeds and plant growth promoting activity of ACC deaminase producing *Bacillus subtilis* (HYT-12-1) on tomato seedlings. World J Microbiol Biotechnol 30(3):835–845
68. Zhang L, Li M, Li Q, Chen C, Qu M, Li M, Wang Y, Shen X (2018a) The catabolite repressor/activator Cra is a bridge connecting carbon metabolism and host colonization in the plant drought resistance-promoting bacterium *Pantoeaalhagi* LTyr-11Z. Appl Environ Microbiol 84(13):e00054–e00018
69. Zhang J, Zhang C, Yang J, Zhang R, Gao J, Zhao X, Zhao J, Zhao D, Zhang X (2018b) Insights into endophytic bacterial community structures of seeds among various *Oryza sativa* L. rice genotypes. J Plant Growth Regul 38(1):1–10
70. Zhang L, Zhang W, Li Q, Cui R, Wang Z, Wang Y, Zhang YZ, Ding W, Shen X (2020) Deciphering the root endosphere microbiome of the desert plant *Alhagisparsifolia* for drought resistance-promoting bacteria. Appl Environ Microbiol 86(11):e02863–e02819

Tables

Table 1 Statistical table of information on upland rice seed samples

Sample ID	Variety names	Indica/Stem	Early/Medium/Late	Breeding places	Abbreviations
19H011	Sanbangqishiluo	Stem	Late	Dehong Prefecture, Yunnan Province	DH
19H012	Benbanggu	Stem	Medium	Gengma County, Yunnan Province	GM
19H013	Zimi	Stem	Medium	Jinping County, Yunnan Province	JP
19H014	Laozaogu	Indica	-	Jiangcheng County, Yunnan Province	JC
19H015	Xiaohonggu	Stem	Medium	Luchun County, Yunnan Province	LC
19H017	Mowanggunei	Stem	Late	Mengla County, Yunnan Province	ML2
19H018	Jixuenuo	Stem	Medium	Gengma County, Yunnan Province	GM
19H019	Fanhaopi	Indica	Medium	Zhenkang County, Yunnan Province	ZK
19H022	Haobuka	Stem	Late	Jinghong Country, Yunnan Province	JH
19H024	Haobayong 1	Stem	Medium	Menglian County, Yunnan Province	ML1
19H029	Lamujia	Stem	Late	Ximeng County, Yunnan Province	XM
19H032	Haogelao	Stem	-	Menglian	ML1

Table 2 Statistical results of endophytic bacterial alpha-diversity in each sample

Samples	OTUs	ACE	Chao	Shannon	Simpson
19H011	1237	2152.41±283.50	1450.72±138.08	2.49±0.11	0.25±0.03
19H012	548	993.58±109.81	681.90±76.09	1.34±0.09	0.44±0.04
19H013	322	1493.91±194.63	565.28±59.46	0.35±0.03	0.85±0.01
19H014	540	1128.40±144.40	824.16±157.67	1.29±0.00	0.40±0.00
19H015	1236	2178.22±93.51	1481.45±145.92	2.11±0.56	0.35±0.15
19H017	1527	2503.33±524.37	1824.46±385.30	2.86±0.49	0.18±0.09
19H018	1281	2187.95±121.99	1547.97±116.65	2.65±0.07	0.18±0.02
19H019	774	1274.84±101.17	830.52±119.33	1.50±0.07	0.49±0.02
19H022	752	1371.88±160.68	895.02±38.10	1.54±0.08	0.44±0.04
19H024	1488	2636.47±208.10	1758.92±99.96	2.27±0.14	0.34±0.04
19H029	706	1301.53±22.02	882.11±40.84	1.65±0.00	0.38±0.00
19H032	860	1544.71±195.49	1101.41±119.54	1.67±0.07	0.35±0.01

Note: The difference of bacterial alpha-diversity based on a 16S rDNA sequence assignment dataset with a 97% sequence similarity threshold in upland rice seeds

Table 3 Statistical results of dominant genera in each upland rice seed sample

Taxonomy	19H011	19H012	19H013	19H014	19H015	19H017
<i>Pantoea</i>	50.70%	78.83%	99.24%	90.28%	63.48%	9.75%
<i>Pseudomonas</i>	4.11%	13.56%	0.11%	1.64%	3.11%	37.24%
<i>Curtobacterium</i>	3.91%	0.22%	0.01%	2.64%	1.18%	19.16%
<i>Microbacterium</i>	14.95%	0.37%	0.01%	2.12%	14.49%	6.05%
<i>Methylobacterium</i>	2.51%	0.19%	0.01%	0.24%	1.93%	5.86%
<i>Agrobacterium</i>	4.53%	0.23%	0.11%	0.10%	2.03%	0.52%
<i>Sphingomonas</i>	0.77%	0.24%	0.04%	0.05%	1.56%	0.96%
<i>Aurantimonas</i>	0.69%	0.22%	0.01%	0.07%	1.06%	0.75%
<i>Rhodococcus</i>	0.07%	0.04%	0.01%	0.01%	0.10%	1.09%
<i>Acidovorax</i>	0.74%	0.02%	0.01%	0.01%	0.23%	0.03%
Taxonomy	19H018	19H019	19H022	19H024	19H029	19H032
<i>Pantoea</i>	30.49%	74.82%	73.06%	68.47%	67.13%	83.78%
<i>Pseudomonas</i>	33.47%	2.84%	14.23%	2.40%	5.21%	5.14%
<i>Curtobacterium</i>	5.05%	13.95%	2.11%	1.50%	19.90%	0.22%
<i>Microbacterium</i>	7.05%	1.21%	0.75%	7.26%	0.36%	0.50%
<i>Methylobacterium</i>	2.17%	0.24%	0.25%	3.34%	0.32%	0.44%
<i>Agrobacterium</i>	1.47%	0.07%	0.22%	2.29%	0.01%	0.61%
<i>Sphingomonas</i>	0.84%	0.17%	0.25%	1.48%	0.11%	0.50%
<i>Aurantimonas</i>	1.17%	0.08%	0.18%	1.45%	0.11%	0.42%
<i>Rhodococcus</i>	0.29%	0.06%	0.06%	0.04%	0.08%	0.01%
<i>Acidovorax</i>	0.03%	0.01%	0.01%	0.12%	0.01%	0

Table 4 Environmental data statistics of origin areas of upland rice

Breeding places	Average annual precipitation /mm	Average annual temperature / °C	Altitude /m
DH	1500	19.2	800-2000
GM	1300	18.8	450-3323
JP	2400	22	1470-5958
JC	2300	18.7	317-2207
LC	2400	16.6	320-2637
ML2	1700	21.8	480-2023
ZK	1600	18.7	510-2468
JH	1400	20.3	485-2196.8
ML1	1400	19.6	497-2603
XM	2800	15.2	590-2458.9

Figures

Figure 1

a, Schematic diagram of geographical location distribution of origin areas. The abbreviations of geographic regions are provided in Table 1. b, Climatic information of origin areas of upland rice seed samples. The letter P stands for annual average precipitation, T for annual average temperature, and A for average altitude. The topographic map of Yunnan Province of China was downloaded from Shuijingzhu Maps Downloader (www.rivermap.cn). Note: The designations employed and the

presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

Figure 2

Relative abundance of shared/unshared phyla in each upland rice seed sample. The abscissa represents the sample name, the ordinate represents the relative abundance of species, each color represents one species, and the corresponding rectangular height represents the relative abundance of species. When judging the relative abundance of a species in a sample, you only need to look at the length of the color rectangle and do not need to accumulate the heights of other colors below the rectangle.

Figure 3

Relative abundance of shared/unshared genus in each upland rice seed sample. The abscissa represents the sample name, the ordinate represents the relative abundance of species, each color represents one species, and the corresponding rectangular height represents the relative abundance of species. When judging the relative abundance of a species in a sample, you only need to look at the length of the color rectangle and do not need to accumulate the heights of other colors below the rectangle.

Figure 4

Classification of samples at the level of 97% sequence similarity (genus, top10). The horizontal represents the name of the sample, and the vertical represents the name of the endophytic bacteria in the sample. In the figure, the color gradients corresponding to relative abundance from 0 to 1 are white, light blue and dark blue. The darker the blue, the higher the relative abundance of species.

Figure 5

Principal Co-ordinates Analysis (PCoA). Ecological differences (two-dimensional) between the different groups and samples in the case of mixed samples at OTU = 0.03. The abscissa and ordinate represent the contribution rate of the principal components 1 and 2 to the distribution of the samples. Each point in the figure represents a sample, and the points of the same color come from the same group.

Figure 6

Non-metric multidimensional scaling (NMDS). Each point represents a sample, and the closer the distance between the two points indicates the smaller the difference of community composition between the two samples.

Figure 7

RDA analysis of endophytic dominant bacteria and environmental factors of upland rice seed samples. The environmental factors are indicated by arrows, and the length of the connection between the arrow and the origin represents the degree of correlation between the environmental factors and the community distribution and species distribution. The longer the length, the greater the correlation, and vice versa. The influence of environmental factors on bacteria is reflected by the cosine of the angle, the smaller the angle, the higher the correlation, and vice versa.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementarydataR1.docx](#)